

Multisensory Research (2023) DOI:10.1163/22134808-bja10102



Does Task-Irrelevant Brightness Modulation Affect Auditory Contrast Processing? Exploring the Interplay Between Temporal Synchrony and Stimulus Salience

Hiu Mei Chow^{1,2,*}, Danielle Briggs^{2,*} and Vivian M. Ciaramitaro^{2,**}

 ¹ Department of Psychology, St. Thomas University, Fredericton, NB, E3B 5G3, Canada
² Department of Psychology, University of Massachusetts Boston, Boston, MA 02126, USA *Authors contributed equally to this article.
**Corresponding author; e-mail: vivian.ciaramitaro@umb.edu ORCID iD: Ciaramitaro: 0000-0001-6767-6441

Received 10 February 2022; accepted 18 May 2023; published online 31 May 2023

Abstract

Stimulus factors such as timing, spatial location, and stimulus effectiveness affect whether and how information across the senses is integrated. Extending recent work highlighting interactions between stimulus factors, here we investigated the influence of visual information on auditory processing, complementing previous studies on the influence of auditory information on visual processing. We hypothesized that task-irrelevant and spatially non-informative visual information would enhance auditory contrast processing, when visual information was at an optimal salience level and changed synchronously with the sound. We asked human observers to indicate the location of an amplitudemodulated white-noise sound, while its loudness against a constant white-noise background varied across trials. To test for the influence of task-irrelevant visual information, we modulated screen brightness smoothly (Experiment 1) or transiently (Experiment 2) in phase or out of phase with the amplitude modulation of the target sound. In addition, to test for the interaction between temporal synchrony and stimulus salience, maximum brightness varied systematically across trials. Auditory contrast thresholds were compared across conditions. Results showed that task-irrelevant visual information did not alter auditory contrast thresholds regardless of the nature of modulation of brightness, contrary to our expectations. Nonetheless, task-irrelevant visual information modulated in phase with the target sound reduced auditory contrast thresholds if we accounted for individual differences in the optimal salience required for the largest multisensory effects. Our results are discussed in light of several stimulus factors that might be critical in modulating multisensory enhancement.

Keywords

auditory contrast, multisensory perception, stimulus effectiveness, temporal coincidence

Published with license by Koninklijke Brill NV | DOI: 10.1163/22134808-bja10102 © HIU MEI CHOW ET AL., 2023 | ISSN: 2213-4794 (print) 2213-4808 (online)

1. Introduction

Integration of information arising from our different senses, multisensory integration, can yield performance benefits in daily activities and laboratory tasks. Multisensory enhancement in performance can be useful in revealing the underlying principles that govern multisensory integration. For example, taskirrelevant, non-auditory information has been demonstrated to enhance auditory processing (Child and Wendt, 1938; Gillmeister and Eimer, 2007; Gregg and Brogden, 1952; Lovelace et al., 2003; Maddox et al., 2015; Odgaard et al., 2004; Thorne and Debener, 2008). A classic study in this line of research demonstrated how a task-irrelevant light enhances one's ability to detect a sound compared to when the sound is presented alone (Lovelace et al., 2003). Maddox et al. (2015) observed improved performance when competing auditory streams included an auditory target that was temporally correlated with a visual stimulus. This study, among others, suggests that multisensory enhancement can arise from temporal cues provided by another sense, allowing one to parse one sound source from other sound sources more easily. Such temporal synchrony allows information to be 'bound' together, or processed as the same event. Besides enhancing performance, multisensory stimuli can also induce suppression in neural activity (e.g., Meredith and Stein, 1986) as well as impair behavioral performance (e.g., Hidaka and Ide, 2015; Ide and Hidaka, 2013). For example, Hidaka and Ide (2015) showed that an auditory white-noise burst presented via headphones impaired orientation discrimination performance of a visual Gabor pattern, when the sound was spatially and temporally aligned with the visual stimulus. Understanding when and how different principles of multisensory integration interact with each other, and across different senses, is fundamental to the understanding of multisensory processing.

In addition to temporal synchrony, other factors contribute to multisensory enhancement, such as the intensity, or effectiveness, of stimuli. Signals from one stimulus can not only facilitate detection of another stimulus in a different modality, but also increase intensity ratings of another stimulus (Gillmeister and Eimer, 2007). Weaker intensities have been shown to yield the largest multisensory enhancements (Meredith and Stein, 1986; Sumby and Pollack, 1954). Thus, when either signal alone is weak, the greatest benefits, or multisensory gains, can be seen when these weak signals are combined, unlike when a single strong signal already provides enough information for effective processing.

Many early studies considered several factors important for multisensory integration, but each in isolation. Yet, in the real world these factors rarely occur alone and recent psychophysical studies have shown how factors interact. Temporal synchrony and stimulus effectiveness are two fundamental properties that interact to influence multisensory processing. For example, Fister *et* *al.* (2016) found that perceived timing (synchronous *vs* asynchronous) of a multisensory event could be influenced by the interaction between the temporal features and the salience of the stimuli. In their experiment, reaction times were faster in synchronous *vs* asynchronous presentations, but findings were more prominent with the more salient stimuli, demonstrating a working relationship between these two multisensory properties. Spatial features can also interact with stimulus effectiveness to modulate the strength of multisensory integration. Multisensory enhancements in localizing a target tend to be greater when stimuli are positioned more peripherally and are presented at lower intensities (Nidiffer *et al.*, 2016).

More recently, Chow et al. (2020) found that multisensory enhancement by an auxiliary sense also shows such interaction. This conclusion is based on the finding that visual contrast processing benefits from a task-irrelevant sound modulated in phase with the visual target, but only when the sound is at an optimal loudness for the individual observer. Taken together, there is strong evidence that stimulus properties, such as synchrony and salience, are interdependent and alter multisensory integration. Properties that interact with the effectiveness of a stimulus can therefore determine the degree of multisensory integration and even the direction of multisensory interactions, that is, whether there will be facilitation or suppression. Cross-modal suppression has been demonstrated via the Colavita visual dominance effect, which describes how two sensory modalities compete with each other, such that vision can suppress the other senses (Colavita, 1974; Hartcher-O'Brien et al., 2008). Sinnet et al. (2007) demonstrated both directions in their experiment that kept the task and multisensory event the same between trials; they found auditory facilitation when subjects responded to the visual stimulus and visual inhibition when subjects responded to the auditory stimulus. Nonetheless, it remains to be addressed whether the interaction between stimulus factors is generalizable to both directions of multisensory interaction, when vision influences audition vs when audition influences vision.

In this study, we investigated the influence of task-irrelevant visual information on auditory processing, complementary to our previous study where we investigated the influence of task-irrelevant auditory information on visual processing (Chow *et al.*, 2020). In order to examine the effects when changes to the visual stimuli are continuous *vs* transient, two types of visual brightness modulation were used in two separate experiments: a sinusoidal modulation was used in Experiment 1, and a square-wave modulation was used in Experiment 2. We hypothesized that the multisensory enhancement induced by an auxiliary sense would be bidirectional and symmetric between vision and audition, such that we would observe similar results as those shown in Chow *et al.* (2020). Specifically, we hypothesized that a task-irrelevant flash would improve the threshold of localizing a sound, but only when the flash was modulated in synchrony with the sound *and* when the flash was presented at an optimal salience level. However, if symmetry between vision and audition cannot be assumed, we expect to see visual suppression of auditory processing, based on the Colavita visual dominance effect (Sinnet *et al.*, 2008). We also hypothesized that such influence of visual information on auditory processing would be more robust when the visual stimulus was transient in Experiment 2 *vs* continuous in Experiment 1. This hypothesis is based on previous findings suggesting that transient events (a square-wave modulation of energy) facilitate multisensory processing relative to continuous events (a sine-wave modulation of energy; Van der Burg *et al.*, 2010).

2. General Materials and Method

2.1. Participants

Altogether, 45 students from the University of Massachusetts, Boston (UMB) participated in the study, 22 and 23 in each experiment. Data from seven participants were excluded from analysis due to poor baseline performance [2 standard deviations (SD) above the group mean (M)] (Experiment 1: n = 1; Experiment 2: n = 3), incomplete dataset (Experiment 1: n = 1; Experiment 2: n = 1), or inattention during the study (Experiment 1: n = 1; Experiment 2: n = 0), yielding a total sample of N = 19 (age: M = 22.3 years, SD = 2.7 years; 16 female) and N = 19 (age: M = 25.0 years, SD = 6.5 years; 16 female) in Experiment 1 and 2, respectively. According to an a priori power analysis performed by G*Power (version 3.1.9.2; Faul et al., 2007), this sample size should be sufficient to support a medium effect size (f = 0.3) at a power of 0.85 in a repeated-measures Analysis of Variance (ANOVA) design. All participants provided written informed consent before participation. They had normal or corrected-to-normal vision and did not report hearing problems. Upon completing the study, participants received US\$10 for every hour of participation or extra credit for an undergraduate course. This experimental protocol was approved by the UMB Institutional Review Board.

2.2. Apparatus and Stimuli

In a sound-proof chamber, an LCD monitor (Tobii TX300; refresh rate = 60 Hz) was positioned 60 cm from the participants; a pair of speakers (JA Audio B3-HTPACK) were placed 55 cm behind the monitor and 35 cm away from the monitor center, and as a result, 33.9° apart from each other (Fig. 1A). A video-based eye tracker (Tobii TX300) was attached to the LCD monitor to record participants' eye positions at 300 Hz.

Auditory and visual stimuli were generated using MATLAB R2014b and Psychtoolbox (version: 3.0.12 beta; Brainard, 1997; Kleiner *et al.*, 2007; Pelli,



Figure 1. (A) Experimental setup illustrating an auditory target displayed *via* the speaker's right-hand side. (B) An illustration of how the auditory contrast of the sound and the brightness of the visual stimulus changed as a function of time: the auditory contrast (amplitude) of the sound was modulated with time at 1 Hz from zero contrast to different maximum contrast levels indicated by black/gray lines. In the BL condition (solid), no flash was presented. In the IP condition (densely dashed), the visual stimulus's brightness was modulated in phase with auditory contrast *via* a sine wave (Experiment 1, top) and a square wave (Experiment 2, bottom). In the OP condition (loosely dashed), the visual stimulus's brightness was modulated 180° out of phase with auditory contrast *via* a sine wave (Experiment 1, top) and a square wave (Experiment 2, bottom). BL, baseline; IP, in phase; OP, out of phase.

1997) and lasted for a maximum of 20 s. The auditory stimuli consisted of two components: a white-noise sound modulated sinusoidally in amplitude at 1 Hz (the signal noise) and a constant white-noise sound (the background noise). The background noise was presented at 34 dB, as measured at the participant's head position. The signal noise (target) was presented *via* one of the two speakers, whereas the background noise was always presented through both speakers. The signal noise started soft, then reached a maximum amplitude (Fig. 1B, lines with different gray shades, also termed the *auditory contrast*), which varied from trial to trial: 30, 30.5, 31, 31.5, 32, 32.5, 33 dB measured at the participant's head position. These values for *auditory contrast* were chosen to provide a range of performance levels from chance (at the lowest levels) to maximum accuracy (at the highest levels). The lower the auditory

contrast, the more challenging for participants to localize the sound source. To account for individual differences in the effective stimulus range, the experimenter adjusted the auditory contrast range after a practice block (see section 2.3. *Procedure*). Fourteen participants performed the task at a lower auditory contrast range (30–32 dB), two at the medium range (30.5–32.5 dB), and four at the higher range (31–33 dB).

The visual stimulus (flash) was a rectangle subtending the entire screen (width = 46.1° ; height = 27.2°). This flash was modulated uniformly in brightness, sinusoidally at 1 Hz in Experiment 1 (Fig. 1B, lines with different gray shades and different dashing), or in a square wave at 1 Hz in Experiment 2 (Fig. 1B), from 60 cd/m² to one of four maximum brightness levels (*visual salience*: 75, 90, 120, 180 cd/m²). The *temporal synchrony* (in phase or out of phase) is related to the temporal offset between the auditory loudness modulation and visual brightness modulation. When the flash started dim at 60 cd/m², the flash modulation was in phase with the loudness modulation (Fig. 1B, densely dashed). By contrast, when the flash started at the maximum brightness level, the flash modulation was out of phase with the loudness modulation (Fig. 1B, loosely dashed). *Visual salience* and *temporal synchrony* of the flash were varied from trial to trial pseudo-randomly. Visual brightness (luminance change) was manipulated as it has been shown to be an effective cue that drives multisensory responses behaviorally (Nidiffer *et al.*, 2018).

2.3. Procedure

Participants completed the experiment in two sessions of one hour each, yielding a total of 900 trials. These 900 trials represented nine visual conditions, including two levels of *temporal synchrony*, four levels of *visual salience*, plus the no-flash baseline condition. Participants completed 20 trials per auditory contrast \times 5 auditory contrast levels for each visual condition, yielding 100 trials per condition and 900 trials in total. To ensure accurate eye-tracking for response collection, participants took a short break and completed a five-point calibration procedure for the eye tracker every 90 trials.

Each trial began with a fixation point at the screen center, which varied in shape, color, size (width = height = $0.9-2.9^{\circ}$), and orientation every 150 ms to capture participants' attention at the center. After the participant fixated central fixation (within a window of width = height = 2.5°) for at least one sample (3.33 ms), the auditory target was presented *via* either the left or right speaker in one of the three visual conditions: in-phase (IP) visual stimulus, out-of-phase (OP) visual stimulus, or no flash (baseline, BL). Participants indicated the location of the auditory target by looking at it. The response was recorded when the horizontal gaze position first reached beyond 8.1° from the center for either side. As a reinforcer, after the response, a black square (width = height = 5.38°) would be presented on the side where the auditory target had

been presented and a short cartoon (width = height = 5.4°) would play for 1 s for correct responses. If participants failed to respond within 20 s, the trial would be aborted, discarded, and repeated immediately.

At the beginning of each experiment, participants completed a practice block of 30 BL trials for practice and stimulus assessment. This practice block included 10 trials each for a low, medium, and high auditory contrast target, but no visual stimulus. By default, the difficulty level was set to the most difficult, least apparent range. If participants failed to reach an accuracy of 90% or above for the loudest auditory target, the stimulus range was adjusted to be easier.

2.4. Data Analysis

2.4.1. Measures

For each observer and each visual condition, accuracy (percent correct) was computed for each auditory contrast level (Fig. 2, dots). A Weibull function (Fig. 2, solid lines) was used to fit accuracy performance across five auditory contrast levels using MATLAB and Psignifit 4.0 (Schutt *et al.*, 2016), using equation (1) and equation (2):



 $\psi(x;\alpha,\beta,\gamma,\lambda) = \gamma + (1-\lambda-\gamma)F_{\rm w}(x;\alpha,\beta) \tag{1}$

Figure 2. Data from a participant, showing the percent correct of each auditory contrast level (dots), the fitted psychometric functions (solid lines), and the threshold values derived (vertical lines) of three visual conditions. Dot color and line color indicate the visual condition, where there could be no visual stimulus (BL, solid), when a visual stimulus was modulated in phase with the sound (IP, densely dashed), or when a visual stimulus was modulated out of phase with the sound (OP, loosely dashed). In this example, the participant showed the smallest auditory contrast threshold in the IP, then BL, then OP conditions. This participant's threshold difference would be positive across all comparisons (BL–IP, OP–BL, OP–IP). The horizontal position of the dots is jittered to reduce overlap. BL, baseline; IP, in phase; OP, out of phase.

H. M. Chow et al. / Multisensory Research (2023)

$$F_{\rm w}(x;\alpha,\beta) = 1 - \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]$$
(2)

where α determines the threshold, β the slope, γ the guess rate, λ the lapse rate. From each psychometric function, we derived a threshold (Fig. 2, dashed lines), defined as the minimum auditory contrast required for 80% accuracy. To quantify the changes in threshold based on the *temporal synchrony* between the visual and auditory stimuli, we computed a difference measure based on the log-transformed thresholds for each comparison of interest (BL-IP, OP-BL, OP-IP) at each visual salience level. Log-transformation was performed to normalize the data distribution. The sign of the threshold difference measures indicated the direction of the effect: a positive BL-IP threshold difference would indicate that an IP flash improved auditory contrast performance relative to no flash; a positive OP-BL would indicate an OP flash suppressed auditory contrast performance relative to no flash; a positive OP-IP would indicate an effect of temporal synchrony, such that an IP flash improved auditory contrast performance relative to an OP flash. For reference, in humans, a just noticeable difference in loudness, or the smallest change which is detectable, for soft sounds like those used in this experiment (30-40 dB) is 1 dB (Mills, 1960).

2.4.2. Statistical Analysis

Statistical analyses and figures were generated in R Studio (R Core Team, 2015, version 1.1.423). For each threshold difference at each visual salience, we computed M, SD, and standard error (SE) across participants. We analyzed both the sign and the magnitude of the threshold difference measures because we were interested in both the direction and the magnitude of multisensory influence. One-sample *t*-tests, corrected for multiple comparisons by False Discovery Rate, were used to test if the difference measures at each visual salience level were significantly different from zero. This analysis was done using 't.test' and 'p.adjust' functions in R. To evaluate whether the null hypothesis could be accepted, a one-sample Bayes Factor (BF) analysis was conducted using the BayesFactor package in R: a BF smaller than 0.33 indicates that there is sufficient evidence to accept the null hypothesis (threshold difference = 0) and a BF larger than 3 indicates that there is sufficient evidence to accept the alternative hypothesis (threshold difference approx. 0). To evaluate the magnitude of the measures, a one-way repeated-measures ANOVA with visual salience as a factor was performed using the ez package (Lawrence, 2016) for each difference measure. Hedges' g, a variant of Cohen's d, and generalized eta-squared (η_G^2) were used to describe effect size. Analysis scripts, data, and the experiment program script are available on Open Science Framework via this link: https://osf.io/cy5gs/?view_only= fa39d63ddd12434eba90203e6a6133f0.

Furthermore, we performed two exploratory analyses to consider individual differences. First, to account for individual differences in where the largest multisensory effects occurred (Mendez-Balbuena et al., 2018), for each participant, we relabeled each visual salience bin in relationship to the visual salience bin yielding the largest multisensory facilitation (max BL-IP), as in our complementary paper (Chow et al., 2020). The visual salience bin with max BL-IP was labeled optimal salience; the adjacent bins were labeled optimal salience ± 1 bin; the furthest bins were labeled optimal salience ± 2 bins. Data were averaged if two bins were involved. We performed the same analyses above to investigate whether the difference measures were significantly different from zero and whether the difference measures were significantly modulated by visual salience. Second, to explore whether individual differences in auditory contrast sensitivity at baseline might explain the strength of multisensory effects, we conducted a Pearson correlation between threshold at BL and each of the threshold differences at 90 cd/m^2 and at the individually optimal salience level, respectively.

3. Experiment 1: Results (Sine-Wave Visual Modulation)

3.1. A Task-Irrelevant Flash Minimally Affects Auditory Contrast Performance

The individual and average threshold difference for each visual salience level and each comparison of interest across the entire sample are plotted in Fig. 3. Generally, the BL–IP difference trended negative (Fig. 3A) and the OP–BL difference trended positive (Fig. 3B) with OP–IP difference trending toward



Figure 3. Individual and average threshold difference of (A) BL–IP, (B) OP–BL, and (C) OP–IP across all participants (n = 19) in Experiment 1 (sine-wave visual modulation). Bar height indicates the average; error bars indicate the values of $M \pm 1$ SE. Dot type indicates the grouping of participants based on the sign of BL–IP difference at 90 cd/m², indicating the IP enhancement group (filled, BL–IP > 0) and IP suppression group (open: BL–IP < 0). As a reminder, a positive BL–IP threshold difference would indicate that an IP flash improved auditory contrast performance relative to no flash; a positive OP–BL would indicate an OP flash suppressed auditory contrast performance relative to no flash; a positive OP–IP would indicate an effect of temporal synchrony. BL, baseline; IP, in phase; OP, out of phase.

zero (Fig. 3C) across observers at most salience levels. Statistically, threshold difference did not differ significantly from zero for any comparisons of interest at any salience levels (ps > 0.07). Similarly, the effect of visual salience did not affect any threshold difference (ps > 0.17). Bayes factor analysis showed that there was sufficient evidence to accept that threshold differences were zero under some situations, for instance, IP-BL threshold differences at 75 cd/m² (BF = 0.28) and 90 cd/m² (BF = 0.24), OP–IP threshold difference at 75 cd/m² (BF = 0.26), OP-BL threshold difference at 150 cd/m² (BF = 0.24), but not the others (BF > 0.33 and BF < 3). This general pattern contradicts our expected results in two ways. First, a tendency for IP suppression was hinted at (based on data visualization) or absent (based on statistical analvsis), in contrast to the IP facilitation we expected. Second, both the IP and OP flash exerted an influence in the same direction, in contrast to the opposite direction we expected. Taken together, these results and patterns suggest that our visual modulation, whether an IP or OP flash, does little to affect auditory contrast performance. If anything, they negatively affect auditory contrast performance.

3.2. Visual Salience Affects the Magnitude of Multisensory Effects

Not all participants showed the largest multisensory effects at the same level of visual salience: five (28%), six (33%), three (17%), and four (22%) participants showed the largest BL–IP threshold difference at 75, 90, 150, and 180 cd/m² respectively. Considering these individual differences, we defined an optimal salience level for each participant based on their largest BL–IP threshold differences across salience levels. The individual and average threshold differences aligned to this optimal salience are plotted in Fig. 4. Whereas IP enhancement was not significant at the optimal salience level (M = 0.001,



Figure 4. Individual and average threshold difference of three comparisons (A, BL–IP; B, OP–BL; C, OP–IP) aligned to the optimal salience bins across all participants in Experiment 1 (sine-wave visual modulation). Bar height indicates the average; error bars indicate the values of $M \pm 1$ SE. Asterisks indicate the *p*-values of one-sample *t*-tests against zero (*, p < 0.05; ***, p < 0.001). BL, baseline; IP, in phase; OP, out of phase.

SD = 0.001, $t_{17} = 1.94$, p = 0.07, g = 0.62), IP suppression was significant at the adjacent salience levels (Fig. 4A). The BL–IP threshold difference at optimal salience ± 1 bin was -0.002 (SD = 0.001, $t_{17} = 2.91$, p = 0.018, g = 0.93) and that at optimal salience ± 2 bin was -0.002 (SD = 0.002, $t_{17} = 2.82$, p = 0.018, g = 0.90). This salience-dependent effect is also supported by a significant effect of salience on the BL–IP threshold difference $(F_{2,34} = 22.75, p < 0.001, \eta_G^2 = 0.23)$. Post-hoc analysis revealed that BL– IP threshold difference at the optimal salience was larger than that of optimal salience ± 1 bin (p < 0.001) and that of optimal salience ± 2 bins (p < 0.001), with no difference between optimal salience ± 1 bin and ± 2 bins (p = 0.75). This result suggests that IP suppression is dependent on optimal salience, which is specific to an individual.

By contrast, no significant difference was found when comparing OP and BL thresholds at any salience level ($ps \ge 0.27$) or when exploring the effect of salience on the magnitude of the OP–BL difference ($F_{2,34} = 0.21$, p = 0.81, $\eta_{\rm G}^2 < 0.01$; Fig. 4B). Taken together, a significant IP suppression and an insignificant OP effect suggest that there was a difference between IP and OP visual modulation, which is supported by our analyses of OP-IP threshold differences (Fig. 4C). OP-IP threshold difference was significantly above zero at the optimal salience level (M = 0.002, SD = 0.002, $t_{17} = 3.156$, p = 0.017, g = 1.0) indicating an IP advantage relative to OP, which is not observed at other salience bins ($ps \ge 0.39$). This salience-dependent effect was also supported by a significant effect of salience on the OP-IP threshold difference $(F_{2,34} = 4.45, p = 0.021, \eta_G^2 = 0.16)$. Post-hoc analysis revealed that the IP benefit relative to OP is significantly larger at the optimal salience than at either the ± 1 or ± 2 bin ($ps \le 0.028$), with no significant difference between the two, suboptimal, neighboring salience levels (p = 0.74). This result suggests that temporal synchrony matters at the optimal salience level, as revealed by the significant OP-IP threshold difference, but that it is not present in other salience bins. Overall, these results show multisensory effects are dependent on visual salience aligned to each participant's optimal salience level.

3.3. Individuals With Lower Auditory Baseline Performance Show a Larger Multisensory Effect

To explore whether individual differences in baseline auditory performance account for the magnitude of multisensory effects, we plotted threshold differences against the baseline auditory threshold in Fig. 5. A positive correlation was found between baseline auditory performance and BL–IP threshold difference, at both 90 cd/m² ($r_{16} = 0.578$, p = 0.012) and at the optimal salience level ($r_{16} = 0.733$, p < 0.001), indicating a larger IP facilitation effect in participants with poorer baseline auditory performance. Such a relationship with



Figure 5. Threshold differences when visual flash was at 90 cd/m² plotted against the threshold at baseline (A–C) in Experiment 1 (sine-wave visual modulation). Similarly, threshold differences when visual flash was at optimal salience level plotted against the threshold at baseline (D–F). Threshold differences include BL–IP (A, D), OP–BL (B, E), and OP–IP (C, F). Dot type in A–C indicates the grouping of participants based on the sign of BL–IP difference at 90 cd/m² indicating the IP enhancement group (filled, BL–IP > 0) and IP suppression group (open: BL–IP < 0). Black lines are the best-fit regression lines. Asterisks indicate the *p*-values of the Pearson correlation test (*, p < 0.05; ***, p < 0.001). BL, baseline; IP, in phase; OP, out of phase.

auditory baseline performance was not found for OP–BL or OP–IP threshold difference ($ps \ge 0.07$).

4. Experiment 2: Results (Square-Wave Visual Modulation)

One possible explanation of the null effects of visual enhancement in Experiment 1 is that the sine-wave modulation of our visual stimulus lacks abrupt changes and onset transients and thus is not optimal for impacting auditory processing (Van der Burg *et al.*, 2010). To address this possibility, in Experiment 2, we kept the experimental design the same, except visual brightness modulation was determined by a square wave (as opposed to a sine wave), introducing abrupt onset and offset transients. We expected visual enhancement of auditory processing would be more readily observed when visual information changed more abruptly in this experiment.



Figure 6. Individual and average threshold difference of (A) BL–IP, (B) OP–BL, and (C) OP–IP across all participants (n = 19) in Experiment 2 (square-wave visual modulation). Bar height indicates the average; error bars indicate the values of $M \pm 1$ SE. Dot type indicates the grouping of participants based on the sign of BL–IP difference at 90 cd/m², indicating the IP enhancement group (filled, BL–IP > 0) and IP suppression group (open: BL–IP < 0). BL, baseline; IP, in phase; OP, out of phase.

Despite the potential benefits offered by transient changes in visual information (relative to the gradual changes in auditory information), we did not find an enhancement of auditory processing at any visual salience level or for any comparisons of interest. This null result is illustrated by Fig. 6, plotting the individual and average threshold difference across the entire sample. All threshold differences trended zero (Fig. 6A-C) across observers at most salience levels. Statistically, threshold difference did not differ significantly from zero for any comparisons of interest at any salience levels (IP enhancement ps = 0.87; OP suppression $ps \ge 0.093$; IP enhancement relative OP $ps \ge 0.48$). No significant effect of visual salience on threshold difference was found (ps > 0.33). Bayes Factor analysis showed that there was sufficient evidence to accept that threshold difference was zero under some situations, for instance, IP–BL threshold differences at 75 cd/m² (BF = 0.26), 150 cd/m² (BF = 0.24), and 180 cd/m² (BF = 0.30), OP-IP threshold difference at 75 cd/m² (BF = 0.25), OP-BL threshold difference at 75 cd/m² (BF = 0.25) and 90 cd/m² (BF = 0.24), but not the others (BF > 0.33 and BF < 3). Taken together, these results and patterns suggest that our visual modulation, whether an IP or OP transient flash, does little to affect auditory contrast performance.

In addition to replicating the null effect of Experiment 1, Experiment 2 also replicated two other results in Experiment 1. First, we again observed that not all participants showed the largest multisensory effects at the same salience level: five (26%), seven (37%), four (21%), and three (16%) participants showed the largest BL–IP threshold difference at 75, 90, 150, and 180 cd/m², respectively. Similar to Experiment 1, we defined an optimal salience level for each participant based on their largest BL–IP threshold difference as BL–IP threshold difference observed across salience levels. The individual and average threshold differences aligned to this optimal salience are plotted in Fig. 7. The



Figure 7. Individual and average threshold difference of three comparisons (A, BL–IP; B, OP–BL; C, OP–IP) aligned to the optimal salience bins across all participants in Experiment 2 (square-wave visual modulation). Bar height indicates the average; error bars indicate the values of $M \pm 1$ SE. Asterisks indicate the *p*-values of one-sample *t*-tests against zero (**, p < 0.01; ***, p < 0.001). BL, baseline; IP, in phase; OP, out of phase.

effect of optimal salience levels is significant on IP–BL threshold difference ($F_{2,36} = 15.67$, p < 0.001, $\eta_G^2 = 0.217$) and IP–OP threshold difference ($F_{2,36} = 9.25$, p < 0.001, $\eta_G^2 = 0.197$, but not OP–IP threshold differences ($F_{2,36} = 0.968$, p = 0.390, $\eta_G^2 = 0.032$). Post-hoc comparison revealed that IP facilitation relative to baseline at the optimal salience level is significantly different from zero (M = 0.002, $t_{18} = 4.574$, p < 0.001, g = 1.42) and is larger than that in the adjacent bins (ps < 0.001). Similar results are found for IP facilitation relative to OP, where IP–OP threshold difference at the optimal salience level is significantly different from zero (M = 0.002, $t_{18} = 3.003$, p = 0.023, g = 0.933) and is larger than that in the adjacent bins (ps < 0.001).

Second, we again found a positive correlation between baseline auditory performance and BL–IP threshold difference, at both 90 cd/m² ($r_{17} = 0.538$, p = 0.020) and the optimal salience level ($r_{17} = 0.757$, p < 0.001), indicating a larger IP facilitation effect in participants with poorer baseline auditory performance (Fig. 8). Interestingly, OP–BL threshold difference at 90 cd/m² also became more negative with increased baseline auditory performance ($r_{17} = -0.53$, p = 0.020), in accordance with a larger multisensory effect with worse baseline performance. Such a relationship with auditory baseline performance (p = 0.32) or OP–IP threshold difference ($ps \ge 0.35$).

5. Discussion

Stimulus properties such as timing, distance, and intensity are rarely independent of each other in everyday environments. Here we investigated interactions between two factors related to multisensory integration, namely temporal synchrony and visual salience, to understand how visual information influences



Figure 8. Threshold differences when visual flash was at 90 cd/m² plotted against the threshold at baseline (A–C) in Experiment 2 (square-wave visual modulation). Similarly, threshold differences when visual flash was at optimal salience level plotted against the threshold at baseline (D–F). Threshold differences include BL–IP (A, D), OP–BL (B, E), and OP–IP (C, F). Dot type in A–C indicates the grouping of participants based on the sign of BL–IP difference at 90 cd/m² indicating the IP enhancement group (filled, BL–IP > 0) and IP suppression group (open: BL–IP < 0). Black lines are the best-fit regression lines. Asterisks indicate the *p*-values of the Pearson correlation test (*, *p* < 0.05; ***, *p* < 0.001). BL, baseline; IP, in phase; OP, out of phase.

auditory processing. We found that task-irrelevant, spatially noninformative, visual information did not affect auditory contrast processing for slow (1 Hz) visual modulations, whether the modulation was smooth (Experiment 1) or transient (Experiment 2). We discuss these null findings in light of previous research and stimulus considerations related to multisensory enhancement by an auxiliary sense.

5.1. Comparing to a Complementary Study on Multisensory Enhancement

Whereas bidirectional cross-talk between the two senses has been established through anatomical (e.g., Falchier *et al.*, 2002; Rockland and Ojima, 2003) and functional (e.g., Eckert *et al.*, 2008; Giard and Peronnet, 1999) connectivity between primary sensory areas, behavioral examination of multisensory interactions often reports dominance of one sense over another depending on task and participants' age (e.g., Colavita, 1974; Sinnett *et al.*, 2007; see

Hirst *et al.*, 2018, for a recent meta-analysis). To the best of our knowledge, whether the interaction between different stimulus factors on multisensory response enhancement by an auxiliary sense is bidirectional and symmetric between vision and audition has not been established prior to our investigation. Thus, we specifically designed this experiment to complement a previous study investigating auditory enhancement of visual processing (Chow *et al.*, 2020).

Generally, our studies found little evidence to support multisensory enhancement by task-irrelevant information presented via a different sense, whether for sound influencing vision (Chow et al., 2020) or vision influencing sound (the current study). For both studies, only after accounting for individual differences in the optimal salience required for achieving the largest multisensory effects did we observe a benefit of multisensory enhancement of auxiliary information modulated in phase with target modality. This multisensory benefit of temporal synchrony was maximum at the optimal salience level, and reduced at neighboring salience levels. This result is in alignment with the notion that a sweet spot of stimulus salience drives the largest multisensory facilitation (e.g., Ross et al., 2007). This is in contrast to the common interpretation of the Principle of Inverse Effectiveness (e.g., Meredith and Stein, 1986; Senkowski et al., 2011; Starke et al., 2020), where multisensory facilitation increases with decreasing stimulus salience. This interaction between stimulus salience and temporal coincidence shown in our work expands the generalizability of findings from previous investigations using different tasks/stimuli (Fister et al., 2016; Nidiffer et al., 2016; Stevenson et al., 2012). Furthermore, our work highlights the importance of stimulus salience of the auxiliary information, lending support for the Principle of Congruent Effectiveness (Otto et al., 2013), where maximum multisensory facilitation is observed when the stimuli are matched in effectiveness. Nevertheless, one caveat of the current optimal salience analysis is that the alignment based on the largest multisensory effect from the same dataset could present a problem of circular logic. Future research should identify the optimal salience for each individual observer first.

5.2. Explaining the Lack of Robust Multisensory Enhancement

Given a number of research studies have reported multisensory enhancement by an auxiliary sense (Gillmeister and Eimer, 2007; Hoefer *et al.*, 2013; Lovelace *et al.*, 2003; Maddox *et al.*, 2015; Odgaard *et al.*, 2004), it is surprising to find a lack of robust multisensory enhancement across observers in the current study. Here we explore several explanations for this null effect.

First, the lack of robust multisensory enhancement is unlikely to be contributed by response bias (Odgaard *et al.*, 2003, 2004) or sensory dominance (Welch *et al.*, 1986) used to explain asymmetrical relations between vision

and audition. Odgaard et al. (2003) found that decreasing the probability of stimulus co-occurrence reduced the influence of a sound on perceived brightness, as did controlling for response bias (Odgaard et al., 2003). By contrast, the influence of light on perceived loudness of a sound was not affected by these manipulations (Odgaard et al., 2004). Together, these two studies show that sound-induced enhancement of perceived brightness (mediated by response bias) and light-induced enhancement of perceived loudness (mediated by stimulus-driven processes) do not share the same mechanism. Along a similar vein, the asymmetry between vision and audition can be due to sensory dominance, resulting from the relatively higher spatial acuity of vision and higher temporal resolution of audition (e.g., Welch et al., 1986), altering their relative reliability (e.g., Alais and Burr, 2004). Nevertheless, two key design considerations should have prevented either factors from explaining our null findings: (1) we used a target location discrimination task - free of response bias — to investigate multisensory enhancement, and (2) the spatial and temporal characteristics of stimuli used in our studies are well above their respective discrimination thresholds. In the current study, the two possible sound locations are 33.9° apart, way above the maximum location blur of broadband noise in adults (3.2°, Haustein and Schirmer, 1970, also see review by Blauert, 1996, p. 39). Furthermore the temporal delay between visual and auditory information was either 0 ms (IP) or 500 ms (OP), of which the difference is larger than \sim 70 ms, the visual temporal integration window in adults (e.g., Freschel et al., 2019; Wutz et al., 2016), or ~50 ms, the audio-visual temporal integration window in adults for simple flashes and beeps (e.g., Keetels and Vroomen, 2005; Zampini et al., 2005). Thus, response bias or sensory dominance is unlikely to explain our null findings.

Alternatively to the above factors, the continuous nature of the stimuli might explain inconsistency with previous studies, but is nevertheless refuted by the null finding in Experiment 2. Our study used a continuous brightness modulation of the screen and an amplitude-modulated auditory white-noise stimulus, whereas previous research reporting visual enhancement of auditory processing typically used transient sensory stimuli like transient flashes and beeps (e.g., Lovelace et al., 2003; Odgaard et al., 2004). Additionally, the transient nature of sound has been shown to be critical to the 'pip-and-pop' effect (Van der Burg et al., 2008), where visual search performance is enhanced by a taskirrelevant sound, and only when it is transiently modulated but not when it is sinusoidally modulated (Van der Burg et al., 2010). Inspired by these findings, we conducted Experiment 2 to see if the expected multisensory enhancement effect would emerge with square-wave modulated visual information, which was not the case. This finding suggests that the change of auxiliary stimuli from continuous to transient is not sufficient in supporting a multisensory enhancement in the current design.

To seek further understanding of how our design is different from prior studies, we compared the current study and previous studies using continuous stimuli (see Table 1) and identified that modulation frequency could be a critical factor. All reviewed studies reporting multisensory enhancement by an auxiliary sense used modulations at higher frequencies (e.g., 7 Hz in Maddox *et al.*, 2015; 6 Hz in Nidiffer *et al.*, 2018; 4 Hz in Parise *et al.*, 2013; 2 Hz in Vuong *et al.*, 2019), whereas our study (and the complementary study, Chow *et al.*, 2020) used a slower modulation (i.e., 1 Hz). The lack of multisensory enhancement when modulation is slow could indicate that multisensory enhancement might be frequency-specific. Given that typical auditory modulation in natural speech is in the 2–7 Hz range (e.g., 5 Hz, Oganian and Chang, 2019; 2–7 Hz, Chandrasekaran *et al.*, 2009), it is possible that multisensory enhancement with continuous stimuli is optimal at a higher modulation range. Future studies would need to consider this issue.

Table 1.

A comparison of the current study with recent studies investigating multisensory enhancement using continuous task-irrelevant stimuli

Authors (year)	Visual stimulus	Auditory stimulus	Task	Direction of multisensory influence	Effect
Chow <i>et al.</i> (2020)	Brightness modulation, 1 Hz, sine wave	Amplitude modulation, 1 Hz, sine wave	Visual localization (2AFC)	Aud. \rightarrow Vis.	No effect; enhancement after aligning to optimal salience
Current Study, Experi- ment 1	Brightness modulation, 1 Hz, sine wave	Amplitude modulation, 1 Hz, sine wave	Auditory localization (2AFC)	Vis. \rightarrow Aud.	No effect; enhancement after aligning to optimal salience; suppression in neighboring bins
Current Study, Experi- ment 2	Brightness modulation, 1 Hz, square wave	Amplitude modulation, 1 Hz, sine wave	Auditory localization (2AFC)	Vis. \rightarrow Aud.	No effect; enhancement after aligning to optimal salience; suppression in neighboring bins
Vuong <i>et al.</i> (2019) Experi- ment 1	Size modulation, 2 Hz	Amplitude modulation, 2 Hz	Discrimination of auditory modulation depth	Vis. \rightarrow Aud.	Enhancement

Authors (year)	Visual stimulus	Auditory stimulus	Task	Direction of multisensory influence	Effect
Vuong <i>et al.</i> (2019) Experi- ment 2	Size modulation, 2 Hz (syn- chronous) or 1 Hz (asyn- chronous)	Amplitude modulation, 2 Hz	Discrimination of auditory modulation depth	Vis. \rightarrow Aud.	Enhancement when synchronous
Vuong <i>et al.</i> (2019) Experi- ment 3	Size modulation, 2 Hz	Amplitude modulation, 2 Hz	Discrimination of visual modulation depth	Aud. \rightarrow Vis.	No effect
Maddox <i>et al.</i> (2015)	Size modulation, 7 Hz	Amplitude modulation, 7 Hz	Perturbation detection of one of the two competing streams (auditory selective attention)	$Vis. \rightarrow Aud.$	Enhancement with matched timecourse
Nidiffer <i>et al.</i> (2018)	Brightness modulation, 6 Hz	Amplitude modulation, 6–7 Hz	Detection of near-threshold amplitude modulation	Vis. & Aud. are both task-relevant	Enhancement increases linearly with temporal coherence
Parise <i>et al.</i> (2013)	Train of transient Gaussian blobs, 4 Hz	Train of white-noise clicks, 4 Hz (at three levels of cross- correlation 1, 0.5, 0)	Spatial offset direction discrimination between visual and auditory stimulus	Vis. & Aud. are both task-relevant	Enhanced binding with high cross-correlation, as indicated by reduced conflict detection performance

Rate

estimation

(decision

making)

events, 8.3 Hz, with feedback task-relevant

Table 1.

(Continued)

Raposo et al. Train of

transient

events,

8.3 Hz,

16.6 Hz

(2012)

Train of

transient

16.6 Hz

Correlation

not affect

multisensory benefit (relative to unisensory condition) \rightarrow abstract qualities being extracted independently before binding

(synchronous or

independent) did

Vis. & Aud.

are both

5.3. Neural Mechanisms of Multisensory Enhancement

20

What neural mechanisms might explain the observed multisensory enhancement? The multisensory properties of sensory areas traditionally considered to be unisensory might be at play. For example, early auditory evoked potentials around 50 ms can be modulated by visual and tactile stimuli that occur simultaneously (Giard and Peronnet, 1999; Molholm *et al.*, 2002; Murray *et al.*, 2005). Similarly, early visual evoked potentials can be modulated by auditory stimuli (e.g., Feng *et al.*, 2017; van der Burg *et al.*, 2011). These early neural modulations have been shown to correlate with behavioral enhancement in performance (e.g., Cappe *et al.*, 2012; McDonald *et al.*, 2013; van der Burg *et al.*, 2011), and might explain our current findings as elaborated below.

Enhancement of multisensory stimuli could also be mediated by phase resetting of neuronal oscillations. Visual stimuli have been shown to systematically align the phase of ongoing slow-frequency neuronal oscillations of individual trials in the auditory cortex in macaques (e.g., Kayser et al., 2008; Perrodin et al., 2015) and in humans (Thorne et al., 2011, also see reviews by Bauer et al., 2020; Thorne and Debener, 2014), yielding improvements in performance (Mercier et al., 2015). Likewise, auditory stimuli can also reset the phase of rhythmic activity in visual cortex (Bauer et al., 2021; Lakatos et al., 2009; Romei et al., 2012), influencing subsequent visual perception (Diederich et al., 2014; Fiebelkorn et al., 2011; Romei et al., 2012). In addition, cross-modal phase resetting has been shown in more complex and naturalistic stimuli such as speech (Mégevand et al., 2020) and rhythmic flashes (O'Connell et al., 2020). Given the rhythmic nature of our stimuli (amplitudemodulated sound and brightness-modulated screen), the auxiliary information that is modulated in phase with the target may produce a phase-resetting effect that enhances the neural response to the target. Future work can combine the current behavioral paradigm with EEG or MEG to explore the link between phase resetting, stimulus salience, and temporal synchrony.

Whereas early neural multisensory enhancement has been shown to follow some known determinants of multisensory interaction, e.g., Principle of Inverse Effectiveness (e.g., Senkowski *et al.*, 2011; Starke *et al.*, 2020), the salience of the target and the auxiliary information are rarely independently modulated in neuroimaging studies. For example, Senkowski *et al.* (2011) reported that the enhancement of early ERPs (40–60 ms post stimulus) in response to a bimodal target is present, particularly for low-intensity stimuli. Yet, they always presented stimulus combinations of matching intensity across the senses (e.g., low–low and high–high). More recently, Starke *et al.* (2020) investigated the neural basis for the visual enhancement of auditory sensitivity combining fMRI and EEG. They found that the behavioral benefit was most pronounced when the sound was soft and that this behavioral gain in sensitivity was related to an increase in the BOLD signal in low-level auditory cortex. However, they only manipulated the loudness of the sound, but not the salience of the visual information. Given this prior work, it is difficult to derive how stimulus salience of the auxiliary sense influences multisensory neural responses, suggesting a gap to be filled by future neuroimaging studies.

Mathematical modeling of behavioral data can provide a mechanistic understanding of multisensory computations involved in the brain. Here we propose considerations for future modeling work after reviewing a recent model relevant to the multisensory phenomenon of interest in our study: the amplification of one sense by an auxiliary sense. Billock and Havig (2018) modeled behavioral and neural data using a simple power law, where the amplified response is a power law of the unamplified response, with a compressive exponent (clustering around 0.85) amplifying weaker target signals more than stronger target signals, in accordance with the Principle of Inverse Effectiveness. They showed that this gated power law also works for other combinations of senses (e.g., auditory enhancement of vision and somatosensory enhancement of audition), as well as neural amplification. Our results provide some further considerations building upon this work: first, our results do not readily support multisensory enhancement using continuously modulated signal at a slow rate, suggesting the exponent factor should be allowed to vary depending on the specific stimuli (transient vs continuous; high vs low modulation frequency) and direction of influence (vision on audition vs audition on vision). Second, our results show that the strength of multisensory enhancement varies as a function of stimulus salience of the auxiliary sense, suggesting that the stimulus salience of the auxiliary sense should potentially be incorporated in the model. Complemented by modeling of neural data (e.g., Billock and Havig, 2018; Billock et al., 2021), future modeling will offer incredible insights into how and where in the brain operations are performed for integrating multisensory signals.

6. Conclusion

Our results find little support for auditory enhancement induced by slowly modulating auxiliary visual information. Task-irrelevant and temporally synchronous visual information only improved auditory contrast processing after accounting for individual differences in the optimal salience required for the largest multisensory effects. These effects are similar to those reported in a previous study examining the influence of task-irrelevant sound on visual contrast processing (Chow *et al.*, 2020). Converged across studies, these results highlight the complexity in how different stimulus factors interact to influence

the magnitude and direction of multisensory interactions. The exact mechanism by which such interactions can be accounted for requires further investigation integrating behavioral, neural, and computational approaches.

Acknowledgements

The authors thank a valuable team of undergraduate research apprentices (including Cheryl Li May, Shamana Pei Chin Chan and Mena Beshara) who helped with data collection. UMB Proposal Development Award (V.M.C.), UMB undergraduate research funds (D.B.), and Canadian Institutes of Health Research Postdoc Fellowship and Michael Smith Health Research BC Trainee Award (H.M.C.) held at University of British Columbia, Vancouver, Canada. Conceptualization, H.M.C. and V.M.C.; Methodology, H.M.C. and V.M.C.; Software, H.M.C. and V.M.C.; Formal Analysis, H.M.C., D.B. and V.M.C.; Investigation, H.M.C. and D.B.; Resources, V.M.C.; Data Curation, H.M.C., D.B. and V.M.C.; Writing — Original Draft Preparation, H.M.C. and D.B.; Writing — Review Editing, H.M.C., D.B. and V.M.C.; Visualization, H.M.C. and D.B.; Supervision, V.M.C.; Project Administration, V.M.C.; Funding Acquisition, V.M.C.

Data Availability

Analysis scripts, data, and the experiment program script are available on Open Science Framework *via* this link:

https://osf.io/cy5gs/?view_only=fa39d63ddd12434eba90203e6a6133f0.

References

- Alais, D. and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration, *Curr. Biol.* 14, 257–262. DOI:10.1016/j.cub.2004.01.029.
- Bauer, A.-K. R., Debener, S. and Nobre, A. C. (2020). Synchronisation of neural oscillations and cross-modal influences, *Trends Cogn. Sci.* 24, 481–495. DOI:10.1016/j.tics.2020.03. 003.
- Bauer, A.-K. R., van Ede, F., Quinn, A. J. and Nobre, A. C. (2021). Rhythmic modulation of visual perception by continuous rhythmic auditory stimulation, *J. Neurosci.* 41, 7065–7075. DOI:10.1523/JNEUROSCI.2980-20.2021.
- Billock, V. A. and Havig, P. R. (2018). A simple power law governs many sensory amplifications and multisensory enhancements, *Sci. Rep.* 8, 7645. DOI:10.1038/s41598-018-25973-w.
- Billock, V. A., Kinney, M. J., Schnupp, J. W. H. and Meredith, M. A. (2021). A simple vectorlike law for perceptual information combination is also followed by a class of cortical multisensory bimodal neurons, *iScience* 24, 102527. DOI:10.1016/j.isci.2021.102527.

Blauert, J. (1996). Spatial Hearing: the Psychophysics of Human Sound Localization. MIT Press, Cambridge, MA, USA. DOI:10.7551/mitpress/6391.001.0001.

Brainard, D. H. (1997). The psychophysics toolbox, Spat. Vis. 10, 433-436.

- Cappe, C., Thelen, A., Romei, V., Thut, G. and Murray, M. M. (2012). Looming signals reveal synergistic principles of multisensory integration, *J. Neurosci.* 32, 1171–1182. DOI:10. 1523/JNEUROSCI.5517-11.2012.
- Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A. and Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech, *PLoS Comput. Biol.* 5, e1000436. DOI:10.1371/ journal.pcbi.1000436.
- Child, I. L. and Wendt, G. R. (1938). The temporal course of the influence of visual stimulation upon the auditory threshold, *J. Exp. Psychol.* 23, 109–127. DOI:10.1037/h0059792.
- Chow, H. M., Leviyah, X. and Ciaramitaro, V. M. (2020). Individual differences in multisensory interactions: the influence of temporal phase coherence and auditory salience on visual contrast sensitivity, *Vision* 4, 12. DOI:10.3390/vision4010012.
- Colavita, F. B. (1974). Human sensory dominance, *Percept. Psychophys.* **16**, 409–412. DOI:10. 3758/BF03203962.
- Diederich, A., Schomburg, A. and van Vugt, M. (2014). Fronto-central theta oscillations are related to oscillations in saccadic response times (SRT): an EEG and behavioral data analysis, *PLoS ONE* 9, e112974. DOI:10.1371/journal.pone.0112974.
- Eckert, M. A., Kamdar, N. V., Chang, C. E., Beckmann, C. F., Greicius, M. D. and Menon, V. (2008). A cross-modal system linking primary auditory and visual cortices: evidence from intrinsic fMRI connectivity analysis, *Hum. Brain Mapp.* **29**, 848–857. DOI:10.1002/hbm. 20560.
- Falchier, A., Clavagnier, S., Barone, P. and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex, J. Neurosci. 22, 5749–5759. DOI:10.1523/ JNEUROSCI.22-13-05749.2002.
- Faul, F., Erdfelder, E., Lang, A.-G. and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences, *Behav. Res. Meth.* 39, 175–191. DOI:10.3758/bf03193146.
- Feng, W., Störmer, V. S., Martinez, A., McDonald, J. J. and Hillyard, S. A. (2017). Involuntary orienting of attention to a sound desynchronizes the occipital alpha rhythm and improves visual perception, *NeuroImage* 150, 318–328. DOI:10.1016/j.neuroimage.2017.02.033.
- Fiebelkorn, I. C., Foxe, J. J., Butler, J. S., Mercier, M. R., Snyder, A. C. and Molholm, S. (2011). Ready, set, reset: stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset, *J. Neurosci.* **31**, 9971–9981. DOI:10. 1523/JNEUROSCI.1338-11.2011.
- Fister, J. K., Stevenson, R. A., Nidiffer, A. R., Barnett, Z. P. and Wallace, M. T. (2016). Stimulus intensity modulates multisensory temporal processing, *Neuropsychologia* 88, 92–100. DOI:10.1016/j.neuropsychologia.2016.02.016.
- Freschl, J., Melcher, D., Kaldy, Z. and Blaser, E. (2019). Visual temporal integration windows are adult-like in 5- to 7-year-old children, *J. Vis.* **19**, 5. DOI:10.1167/19.7.5.
- Giard, M. H. and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study, *J. Cogn. Neurosci.* 11, 473–490. DOI:10.1162/089892999563544.
- Gillmeister, H. and Eimer, M. (2007). Tactile enhancement of auditory detection and perceived loudness, *Brain Res.* 1160, 58–68. DOI:10.1016/j.brainres.2007.03.041.
- Gregg, L. W. and Brogden, W. J. (1952). The effect of simultaneous visual stimulation on absolute auditory sensitivity, J. Exp. Psychol. 43, 179–186. DOI:10.1037/h0061877.

- Hartcher-O'Brien, J., Gallace, A., Krings, B., Koppen, C. and Spence, C. (2008). When vision 'extinguishes' touch in neurologically-normal people: extending the Colavita visual dominance effect, *Exp. Brain Res.* 186, 643–658. DOI:10.1007/s00221-008-1272-5.
- Haustein, B. G. and Schirmer, W. (1970). Messeinrichtung zur Untersuchung des Richtungslokalisationsvermögens, *Hochfrequenztech. Elektroakustik* **79**, 96–101.
- Hidaka, S. and Ide, M. (2015). Sound can suppress visual perception, *Sci. Rep.* 5, 10483. DOI:10.1038/srep10483.
- Hirst, R. J., Cragg, L. and Allen, H. A. (2018). Vision dominates audition in adults but not children: a meta-analysis of the Colavita effect, *Neurosci. Biobehav. Rev.* 94, 286–301. DOI:10. 1016/j.neubiorev.2018.07.012.
- Hoefer, M., Tyll, S., Kanowski, M., Brosch, M., Schoenfeld, M. A., Heinze, H.-J. and Noesselt, T. (2013). Tactile stimulation and hemispheric asymmetries modulate auditory perception and neural responses in primary auditory cortex, *NeuroImage* 79, 371–382. DOI:10.1016/j. neuroimage.2013.04.119.
- Ide, M. and Hidaka, S. (2013). Tactile stimulation can suppress visual perception, *Sci. Rep.* **3**, 3453. DOI:10.1038/srep03453.
- Kayser, C., Petkov, C. I. and Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex, *Cereb. Cortex* 18, 1560–1574. DOI:10.1093/cercor/bhm187.
- Keetels, M. and Vroomen, J. (2005). The role of spatial disparity and hemifields in audio-visual temporal order judgments, *Exp. Brain Res.* 167, 635–640. DOI:10.1007/s00221-005-0067-1.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R. and Broussard, C. (2007). What's new in Psychtoolbox-3, *Perception* 36, 1–16.
- Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C. and Schroeder, C. E. (2009). The leading sense: supramodal control of neurophysiological context by attention, *Neuron* 64, 419–430. DOI:10.1016/j.neuron.2009.10.014.
- Lawrence, M. A. (2016). ez: easy analysis and visualization of factorial experiments [Computer Software Manual]. Retrieved from https://cran.r-project.org/web/packages/ez/index.html (R package version 4).
- Lovelace, C. T., Stein, B. E. and Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection, *Cogn. Brain Res.* 17, 447–453. DOI:10.1016/S0926-6410(03)00160-5.
- Maddox, R. K., Atilgan, H., Bizley, J. K. and Lee, A. K. C. (2015). Auditory selective attention is enhanced by a task-irrelevant temporally coherent visual stimulus in human listeners, *eLife* 4, e04995. DOI:10.7554/eLife.04995.
- McDonald, J. J., Störmer, V. S., Martinez, A., Feng, W. and Hillyard, S. A. (2013). Salient sounds activate human visual cortex automatically, *J. Neurosci.* 33, 9194–9201. DOI:10. 1523/JNEUROSCI.5902-12.2013.
- Mégevand, P., Mercier, M. R., Groppe, D. M., Golumbic, E. Z., Mesgarani, N., Beauchamp, M. S., Schroeder, C. E. and Mehta, A. D. (2020). Crossmodal phase reset and evoked responses provide complementary mechanisms for the influence of visual speech in auditory cortex, *J. Neurosci.* 40, 8530–8542. DOI:10.1523/JNEUROSCI.0555-20.2020.
- Mendez-Balbuena, I., Arrieta, P., Huidobro, N., Flores, A., Lemuz-Lopez, R., Trenado, C. and Manjarrez, E. (2018). Augmenting EEG-global-coherence with auditory and visual noise, *Medicine* 97, e12008. DOI:10.1097/MD.000000000012008.

- Mercier, M. R., Molholm, S., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H. and Foxe, J. J. (2015). Neuro-oscillatory phase alignment drives speeded multisensory response times: an electro-corticographic investigation, *J. Neurosci.* 35, 8546–8557. DOI:10.1523/ JNEUROSCI.4527-14.2015.
- Meredith, M. A. and Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration, *J. Neurophysiol.* 56, 640–662. DOI:10.1152/jn.1986.56.3.640.
- Mills, A. W. (1960). Lateralization of high-frequency tones, J. Acoust. Soc. Am. **32**, 132–134. DOI:10.1121/1.1907864.
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E. and Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study, *Cogn. Brain Res.* 14, 115–128. DOI:10.1016/s0926-6410(02)00066-6.
- Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., Schroeder, C. E. and Foxe, J. J. (2005). Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment, *Cereb. Cortex* 15, 963–974. DOI:10.1093/cercor/bhh197.
- Nidiffer, A. R., Stevenson, R. A., Fister, J. K., Barnett, Z. P. and Wallace, M. T. (2016). Interactions between space and effectiveness in human multisensory performance, *Neuropsychologia* 88, 83–91. DOI:10.1016/j.neuropsychologia.2016.01.031.
- Nidiffer, A. R., Diederich, A., Ramachandran, R. and Wallace, M. T. (2018). Multisensory perception reflects individual differences in processing temporal correlations, *Sci. Rep.* 8, 14483. DOI:10.1038/s41598-018-32673-y.
- O'Connell, M. N., Barczak, A., McGinnis, T., Mackin, K., Mowery, T., Schroeder, C. E. and Lakatos, P. (2020). The role of motor and environmental visual rhythms in structuring auditory cortical excitability, *iScience* 23, 101374. DOI:10.1016/j.isci.2020.101374.
- Odgaard, E. C., Arieh, Y. and Marks, L. E. (2003). Cross-modal enhancement of perceived brightness: sensory interaction versus response bias, *Percept. Psychophys.* 65, 123–132. DOI:10.3758/BF03194789.
- Odgaard, E. C., Arieh, Y. and Marks, L. E. (2004). Brighter noise: sensory enhancement of perceived loudness by concurrent visual stimulation, *Cogn. Affect. Behav. Neurosci.* 4, 127– 132. DOI:10.3758/CABN.4.2.127.
- Oganian, Y. and Chang, E. F. (2019). A speech envelope landmark for syllable encoding in human superior temporal gyrus, *Sci. Adv.* **5**, eaay6279. DOI:10.1126/sciadv.aay6279.
- Otto, T. U., Dassy, B. and Mamassian, P. (2013). Principles of multisensory behavior, J. Neurosci. 33, 7463–7474. DOI:10.1523/JNEUROSCI.4678-12.2013.
- Parise, C. V., Harrar, V., Ernst, M. O. and Spence, C. (2013). Cross-correlation between auditory and visual signals promotes multisensory integration, *Multisens. Res.* 26, 307–316. DOI:10. 1163/22134808-00002417.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies, *Spat. Vis.* 10, 437–442.
- Perrodin, C., Kayser, C., Logothetis, N. K. and Petkov, C. I. (2015). Natural asynchronies in audiovisual communication signals regulate neuronal multisensory interactions in voicesensitive cortex, *Proc. Natl Acad. Sci. USA* 112, 273–278. DOI:10.1073/pnas.1412817112.

- R Core Team (2015). R: a Language and Environment for Statistical Computing [Computer Software Manual]. Vienna, Austria. Retrieved from http://www.R-project.org/.
- Raposo, D., Sheppard, J. P., Schrater, P. R. and Churchland, A. K. (2012). Multisensory decision-making in rats and humans, J. Neurosci. 32, 3726–3735. DOI:10.1523/ JNEUROSCI.4998-11.2012.
- Rockland, K. S. and Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey, *Int. J. Psychophysiol.* 50, 19–26. DOI:10.1016/s0167-8760(03)00121-1.
- Romei, V., Gross, J. and Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception, *Curr. Biol.* 22, 807–813. DOI:10.1016/j.cub.2012.03.025.
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C. and Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments, *Cereb. Cortex* 17, 1147–1153. DOI:10.1093/cercor/bhl024.
- Schütt, H. H., Harmeling, S., Macke, J. H. and Wichmann, F. A. (2016). Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data, *Vis. Res.* 122, 105–123. DOI:10.1016/j.visres.2016.02.002.
- Senkowski, D., Saint-Amour, D., Höfle, M. and Foxe, J. J. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness, *NeuroImage* 56, 2200–2208. DOI:10.1016/j.neuroimage.2011.03.075.
- Sinnett, S., Spence, C. and Soto-Faraco, S. (2007). Visual dominance and attention: the Colavita effect revisited, *Percept. Psychophys.* **69**, 673–686. DOI:10.3758/bf03193770.
- Starke, J., Ball, F., Heinze, H.-J. and Noesselt, T. (2020). The spatio-temporal profile of multisensory integration, *Eur. J. Neurosci.* 51, 1210–1223. DOI:10.1111/ejn.13753.
- Stevenson, R. A., Krueger Fister, J. K., Barnett, Z. P., Nidiffer, A. R. and Wallace, M. T. (2012). Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance, *Exp. Brain Res.* 219, 121–137. DOI:10.1007/s00221-012-3072-1.
- Sumby, W. H. and Pollack, I. (1954). Visual contribution to speech intelligibility in noise, J. Acoust. Soc. Am. 26, 212–215. DOI:10.1121/1.1907309.
- Thorne, J. D. and Debener, S. (2008). Irrelevant visual stimuli improve auditory task performance, *NeuroReport* 19, 553–557. DOI:10.1097/WNR.0b013e3282f8b1b6.
- Thorne, J. D. and Debener, S. (2014). Look now and hear what's coming: on the functional role of cross-modal phase reset, *Hear. Res.* 307, 144–152. DOI:10.1016/j.heares.2013.07.002.
- Thorne, J. D., De Vos, M., Viola, F. C. and Debener, S. (2011). Cross-modal phase reset predicts auditory task performance in humans, *J. Neurosci.* **31**, 3853–3861. DOI:10.1523/JNEUROSCI.6176-10.2011.
- Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W. and Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search, *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1053–1065. DOI:10.1037/0096-1523.34.5.1053.
- Van der Burg, E., Cass, J., Olivers, C. N. L., Theeuwes, J. and Alais, D. (2010). Efficient visual search from synchronized auditory signals requires transient audiovisual events, *PLoS ONE* 5, e10664. DOI:10.1371/journal.pone.0010664.
- Van der Burg, E., Talsma, D., Olivers, C. N. L., Hickey, C. and Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects, *NeuroImage* 55, 1208–1218. DOI:10.1016/j.neuroimage.2010.12.068.

- Vuong, Q. C., Laing, M., Prabhu, A., Tung, H. I. and Rees, A. (2019). Modulated stimuli demonstrate asymmetric interactions between hearing and vision, *Sci. Rep.* 9, 7605. DOI:10.1038/ s41598-019-44079-5.
- Welch, R. B., DuttonHurt, L. D. and Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception, *Percept. Psychophys.* 39, 294–300. DOI:10.3758/bf03204939.
- Wutz, A., Muschter, E., van Koningsbruggen, M. G., Weisz, N. and Melcher, D. (2016). Temporal integration windows in neural processing and perception aligned to saccadic eye movements, *Curr. Biol.* 26, 1659–1668. DOI:10.1016/j.cub.2016.04.070.
- Zampini, M., Guest, S., Shore, D. I. and Spence, C. (2005). Audio-visual simultaneity judgments, *Percept. Psychophys.* 67, 531–544. DOI:10.3758/BF03193329.